

Ben Golan

Research Engineer — LLM Post-Training — Reinforcement Learning — Personalization
Tel Aviv, Israel — +972 54 5635563 — golanben1@gmail.com
linkedin.com/in/ben-golan-4b1b994

Summary

Research Engineer with 15+ years of experience in machine learning, reinforcement learning, and multimodal systems. Currently contributing to large language model (LLM) post-training and evaluation at Google Gemini, with focus on reward modeling, alignment, factuality, and robustness. Experienced in designing experimental frameworks, reinforcement learning environments, synthetic data curation, and scalable evaluation systems. Interested in advancing post-training methodologies and personalization strategies for frontier foundation models.

Core Expertise

- LLM Post-Training: RLHF, reward modeling, critique-based learning
- Reinforcement Learning: policy gradients, DQN, environment design
- Personalization Adaptation: user-conditioned generation, evaluation design
- Synthetic data curation and benchmark construction
- Experimental design, ablation analysis, statistical testing
- PyTorch-based model experimentation and large-scale evaluation

Experience

Google – Gemini Post-Training

Senior AI Research Engineer

Tel Aviv

Feb 2025 – Present

- Contributed to post-training strategies for large language models, focusing on reward modeling, alignment, and hallucination mitigation.
- Co-developed the FACTS Leaderboard benchmark suite (arXiv:2512.10791), evaluating multimodal, parametric, search-based, and grounded factuality of LLMs.
- Designed experimental frameworks for grounded reasoning, long-context behavior, and judge-model-based evaluation.
- Built reusable evaluation pipelines enabling systematic comparison across model checkpoints.
- Investigated reinforcement-learning-driven improvements and critique-based optimization of model outputs.

Sentrycs (Acquired by ONDAS)

Senior Research Engineer

Tel Aviv

2021 – 2024

- Led end-to-end modeling efforts from data collection through experimentation and deployment.

- Designed probabilistic tracking systems and supervised learning pipelines.
- Conducted structured ablation experiments to assess robustness under distribution shifts.

SAIPS (Ford Motor Company)
Senior ML Engineer

Tel Aviv
 2015 – 2019

- Developed deep learning models for perception tasks (object detection, segmentation).
- Implemented reinforcement learning algorithms for autonomous decision-making.
- Designed experimental protocols to compare alternative modeling strategies.

Robin Healthcare
Senior ML Engineer

Tel Aviv
 2019 – 2021

- Built multimodal NLP + ASR pipelines for clinical documentation.
- Fine-tuned transformer-based models for structured dialogue understanding.
- Developed scalable training and evaluation workflows.

Publications

The FACTS Leaderboard: A Comprehensive Benchmark for Large Language Model Factuality.

Aileen Cheng, Alon Jacovi, Amir Globerson, **Ben Golan** (equal contribution), et al.
 arXiv:2512.10791 (cs.CL, cs.AI), 2025.

Comprehensive benchmark suite evaluating factuality across multimodal, parametric, search-based, and grounded LLM settings.

Education

Tel Aviv University

M.Sc. Electrical Engineering (Machine Learning, Communication Systems)

B.Sc. Physics and Electrical Engineering, Magna Cum Laude

PhD-level coursework in Advanced Machine Learning (Weizmann Institute).

Technical Skills

Languages: Python, C++

Frameworks: PyTorch, JAX, TensorFlow

Domains: LLM post-training, reinforcement learning, personalization, multimodal modeling

Practices: Experimental design, statistical evaluation, scalable ML systems